A Novel Gamma Filter for Positive Parameter Estimation

Corey Marcus[1] and Renato Zanetti[2]

***Abstract*— This work develops a new recursive Bayesian estimation algorithm for gamma-distributed random variables. The time prediction and measurement update steps are developed, and both are shown to have analytic closed formed solutions under certain conditions. Furthermore, the update is shown to be unbiased, this is theoretically guaranteed by the choice of estimator and demonstrated numerically through Monte Carlo methods with a simple example.**

I. Introduction

Strictly positive random variables are of interest in many situations such as, mass, temperature (in Kelvin), or length. In the field of simultaneous localization and mapping, authors such as Engel *et al* or Marcus and Zanetti estimate the strictly positive depth of features in camera images [1], [2]. Whatever the application, these variables are generally estimated through some form of Bayesian estimation, or recursive Bayesian estimation for situations with multiple measurements and/or dynamics [3]. In cases with linear and gaussian systems, this estimator takes the form of the Kalman Filter [4]. Real systems are never truely linear and gaussian, however assumptions of linearity and gaussianity result in adequate performance for a large number of problems. A variety of nonlinear extensions of the Kalman Filter such as the Extended Kalman Filter (EKF) [5] have been created in case nonlinearities are large enough to preclude use of a traditional Kalman Filter.

The gaussian distribution has a domain of $(-\infty, \infty)$, making it ill-posed for representing the distribution of strictly positive random variables. Govaers and Alqaderi have previously identified this limitation and developed a novel Bayesian recursion filter for gamma-distributed random variables [6]. The filter they propose is biased, but they also propose a correction that, at the cost of a higher estimation error variance, produces an unbiased estimator. Our previous work [2] developed an update formula when the measurements are functions of the inverse of a Gamma-distributed random variable.

This paper presents a novel recursive Bayesian filter for Gamma-distributed random variables. The choice of estimate is the one satisfying the Minimum Mean Square Error criteria [3] which assures its unbiasedness. A simple numerical example is provided, and Monte Carlo analysis shows that the estimator is unbiased and consistent.

*This work was supported by a NASA Space Technology Graduate Research Opportunity, NASA grant number: 80NSSC20K1195.

[1]Corey Marcus is a Student at the Department Aerospace and Engineering Mechanics The University of Texas at Austin, USA `cmarcus@utexas.edu`

[2]Renato Zanetti is an Assistant Professor at the Department of Aerospace and Engineering Mechanics The University of Texas at Austin, USA

This paper focuses on the estimation of the unknown outcome of a Gamma-distributed random variable via noisy measurements of it. As such, this work differs from the existing literature on estimating the parameters (shape and scale/rate) of a Gamma distribution from several samples drawn from it [7], [8].

II. Formulation of the Problem

We primarily concern ourselves with positive random variables whose probability density functions are represented with the gamma distribution. The gamma distribution for a random variable, x, can be defined by two parameters, shape, α, and rate, β, as shown in Equation (1).

$$\mathscr{G}(x;\alpha,\beta) = \frac{\beta^{\alpha} x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \tag{1}$$

where $\Gamma(\cdot)$ is the gamma function. The mean and variance of x are:

$$E[x] = \frac{\alpha}{\beta} \tag{2}$$

$$\mathrm{var}[x] = \frac{\alpha}{\beta^2} \tag{3}$$

We intend to recursively estimate a dynamic state x_i given measurements y_i which arrive at each discrete time-step t_i. Equation (4) represent the time-evolution of x_i, with the caveat that it always remains positive. The time evolution process may also be corrupted by stochastic noise, v_i, with a known distribution, $p_{v_i}(v_i)$. The typical assumption that x_i and v_i are independent is made.

$$x_{i+1} = f(x_i, v_i) \tag{4}$$

The measurement process, Equation (5), produces a measurement, y_i at each time-step and is corrupted by stochastic noise, w_i, with a known distribution, $p_{w_i}(w_i)$. It is also assumed that x_i and w_i are independent.

$$y_i = h(x_i, w_i) \tag{5}$$

Our goal is to produce a recursive Bayesian estimate of x_i, $\hat{x}_i$. This estimate is a function of the posterior distribution, $p_{x_i}(x_i|Y_i)$, which is the distribution of x_i conditioned on all measurements up to and including the i^{th}. As a shorthand notation, we denote this sequence of measurements Y_i.

III. METHODOLOGY

In general, recursive Bayesian estimation can be formulated exactly without approximations when the distribution of x_i is a conjugate prior of the likelihood function, $L(x_i|y_i) = p_{y_i}(y_i|x_i)$. If $h(\cdot)$ can be inverted such that $w_i = h^{-1}(y_i, x_i)$, the likelihood can be easily found through a transformation of random variables [9] approach as shown in Equation (6).

$$L(x_i|y_i) = \left|\frac{\partial h^{-1}(y_i,x_i)}{\partial y_i}\right| p_{w_i}\left(h^{-1}\left(y_i,x_i\right)\right) \tag{6}$$

Whenever $p_{x_i}(x_i)$ is a conjugate prior to $L(x_i|y_i)$, Bayes Rule, Equation (7), can be applied to produce a gamma-distributed posterior and an exact recursive filter can be formulated.

$$p(x_i|Y_i) = \frac{L(x_i|y_i)p(x_i|Y_{i-1})}{\int L(x_i|y_i)p(x_i|Y_{i-1})dx_i} \tag{7}$$

Our filtering algorithm consists of a prediction step that maps the state forward in time and an update step which incorporates the current measurement. The recursion is obtained by starting the next prediction step with the post-update distribution.

A. Predict

We begin with a gamma distribution representing the distribution of estimate of random variable x_{i-1} conditioned on measurement sequence Y_{i-1}, $p_{x_{i-1}}(x_{i-1}|Y_{i-1}) = \mathscr{G}(x_{i-1};\alpha_{i-1|i-1},\beta_{i-1|i-1})$. The prediction of the distribution at time i may be found through integration as:

$$\begin{aligned} p(x_i|Y_{i-1}) &= \int p(x_i,x_{i-1}|Y_{i-1})\ dx_{i-1} && (8)\\ &= \int p(x_i|x_{i-1},Y_{i-1})\ p(x_{i-1}|Y_{i-1})\ dx_{i-1} && (9)\end{aligned}$$

The above integral is typically not solvable in closed form, unless very specific assumptions on the transitional density, $p(x_i|x_{i-1})$, are made. In this work we concentrate on the following propagation model:

$$x_i = v_{i-1}x_{i-1} \tag{10}$$

where $v_{i-1} > 0$.

v_{i-1} can either be deterministic or a random variable. When it is deterministic the transformed random variable is also Gamma-distributed as $x_{i|i-1} \sim \mathscr{G}\left(x_i;\alpha_{i-1|i-1},\beta_{i-1|i-1}/v_{i-1}\right)$. This can be proven using the moment generating function (MGF) of a gamma-distributed random variable. We provide the MGF of a generic gamma-distributed random variable, y, in Equation (11)

$$\begin{aligned} &y \sim \mathscr{G}\left(\alpha,\beta\right)\\ &M_y(t) = E\left[e^{ty}\right] = \left(\frac{\beta}{\beta-t}\right)^{\alpha} && (11)\end{aligned}$$

The distribution of x_i for deterministic v_{i-1} can be found by inspecting the MGF of x_i.

$$\begin{aligned} M_{x_i}(t) &= M_{v_{i-1}x_{i-1}}(t) && (12)\\ &= E\left[e^{t(v_{i-1}x_{i-1})}\right] = E\left[e^{(tv_{i-1})x_{i-1}}\right] && (13)\\ &= M_{x_{i-1}}(v_{i-1}t) && (14)\\ &= \left(\frac{\beta_{i-1|i-1}}{\beta_{i-1|i-1}-v_{i-1}t}\right)^{\alpha_{i-1|i-1}} && (15)\\ &= \left(\frac{\beta_{i-1|i-1}/v_{i-1}}{\beta_{i-1|i-1}/v_{i-1}-t}\right)^{\alpha_{i-1|i-1}} && (16)\end{aligned}$$

By recalling that the moment generating function uniquely determines the distribution we can conclude that $x_{i|i-1} \sim \mathscr{G}\left(x_i;\alpha_{i-1|i-1},\beta_{i-1|i-1}/v_{i-1}\right)$.

Alternatively, v_i can be a random variable independent from x_i. In this case, $p(x_i|Y_{i-1})$ is not generally a gamma distribution. However, we will approximate it as one through a moment matching procedure. Let $a_i = \mathrm{E}[v_i]$ and define $v_i = a_i + \delta v_i$ where $\delta v_{i-1}x_{i-1}$ is a zero-mean process noise term. The propagation can then be rewritten as

$$x_i = a_{i-1}\ x_{i-1} + \delta v_{i-1}\ x_{i-1} \tag{17}$$

As previously mentioned, the exact value of $p(x_i|Y_{i-1})$ is often not obtainable in closed form and/or not a gamma distribution. We require the mean and variance of x_i in Equation (17) to moment match. The mean and variance of the product of two independent random variables v_{i-1} and x_{i-1} is found according to Equations (18) and (19).

$$E[v_{i-1}x_{i-1}] = E[v_{i-1}]E[x_{i-1}] \tag{18}$$

$$\mathrm{var}[v_{i-1}x_{i-1}] = E[v_{i-1}^2]E[x_{i-1}^2] - \left(E[v_{i-1}]E[x_{i-1}]\right)^2 \tag{19}$$

Once the mean and variance have been found, the parameters $\alpha_{i|i-1}$ and $\beta_{i|i-1}$ of the time-propagated gamma distribution are chosen such that the mean and variance of $p(x_i|Y_{i-1}) = \mathscr{G}(x_i;\alpha_{i|i-1},\beta_{i|i-1})$ match the values calculated with Equations (18) and (19).

Regardless of the method used, we arrive at a gamma distribution which exactly or approximately represents the distribution of x_i conditioned on measurement sequence Y_{i-1}.

$$p(x_i|Y_{i-1}) = \mathscr{G}(x_i;\alpha_{i|i-1},\beta_{i|i-1}) \tag{20}$$

B. Update

We begin with a gamma-distributed state, $x_i \sim \mathscr{G}(x_{i|i-1};\alpha_{i|i-1},\beta_{i|i-1})$, which is conditioned on all previous measurements up to and including y_{i-1}. We assume the state observations are themselves positive quantities and we model them as

$$y_i = w_i x_i \tag{21}$$

where w_i is a random variable independent from x_i and distributed as an inverse gamma, $w_i \sim \mathscr{G}^{-1}(w_i;\alpha_{w_i},\beta_{w_i})$. The inverse gamma distribution is a two parameter distribution family which is closely related to the gamma distribution. If

$w_i \sim \mathscr{G}^{-1}(x_i; \alpha_{w_i}, \beta_{w_i})$, $1/w_i \sim \mathscr{G}(w_i; \alpha_{w_i}, \beta_{w_i})$. The distribution is provided for reference in Equation (22).

$$p_{w_i}(w_i) = \mathscr{G}^{-1}(w_i; \alpha_{w_i}, \beta_{w_i}) = \frac{\beta_{w_i}^{\alpha_{w_i}} w_i^{-\alpha_{w_i}-1} e^{-\beta_{w_i}/w_i}}{\Gamma(\alpha_{w_i})} \quad (22)$$

Our observation model differs from the traditional zero-mean additive noise model. As an aside, we show that the two models are more similar than they may appear. Let $\mathrm{E}[w_i] = b_i$, then the proposed measurement model hence reduces to

$$y_i = b_i \ x_i + x_i \ \delta w_i \quad (23)$$

where $\delta w_i = w_i - \mathrm{E}[w_i]$ is zero-mean. In this alternative form, $x_i \delta w_i$ can be thought of as additive, zero-mean noise. The model in Equation (21) has been chosen for mathematical convenience.

We aim to apply Bayes Rule, Equation (7), to find the posterior. First, we note that $h(\cdot)$ can be inverted as

$$w_i = h^{-1}(x_i, y_i) = \frac{y_i}{x_i} \quad (24)$$

We can now find $L(x_i|y_i)$ with a transformation of random variables approach [9] shown in Equation (6),

$$\begin{aligned} L(x_i|y_i) &= \left| \frac{\partial h^{-1}(y_i,x_i)}{\partial y_i} \right| p_{w_i}\left(h^{-1}(y_i,x_i)\right) \\ &= \left| \frac{\partial}{\partial y_i} \frac{y_i}{x_i} \right| p_{w_i}\left(h^{-1}(y_i,x_i)\right) \\ &= \frac{1}{x_i} p_{w_i}\left(h^{-1}(y_i,x_i)\right) \\ L(x_i|y_i) &= \frac{1}{x_i} \frac{\beta_{w_i}^{\alpha_{w_i}} (y_i/x_i)^{-\alpha_{w_i}-1} e^{-\beta_{w_i} x_i/y_i}}{\Gamma(\alpha_{w_i})}. \end{aligned} \quad (25)$$

Next, we perform a direct multiplication of $L(x_i|y_i)$ and $p(x_{i|i-1})$ to find the numerator in Bayes Rule. Note the denominator is a normalizing constant to ensure the distribution integrates to one and thus does not need explicit calculation. We are able to claim

$$p(x_i|y_i) \propto L(x_i|y_i) p(x_{i|i-1}) \quad (26)$$

$$\propto \frac{\beta_{w_i}^{\alpha_{w_i}} (y_i/x_i)^{-\alpha_{w_i}-1} e^{-\beta_{w_i} x_i/y_i}}{x_i \Gamma(\alpha_{w_i})} \frac{\beta_{i|i-1}^{\alpha_{i|i-1}} x^{\alpha_{i|i-1}-1} e^{-\beta_{i|i-1} x_i}}{\Gamma(\alpha_{i|i-1})} \quad (27)$$

$$\propto x_i^{\alpha_{w_i}+\alpha_{i|i-1}-1} e^{-x_i\left(\frac{\beta_{w_i}}{y_i}+\beta_{i|i-1}\right)} \quad (28)$$

From this we can conclude

$$p(x_i|y_i) = \mathscr{G}(x_{i|i}; \alpha_{i|i}, \beta_{i|i}) \quad (29)$$

where

$$\alpha_{i|i} = \alpha_{w_i} + \alpha_{i|i-1} \quad (30)$$

and

$$\beta_{i|i} = \frac{\beta_{w_i}}{y_i} + \beta_{i|i-1} \quad (31)$$

Having found the posterior distribution $p_{x_i}(x_i|Y_i)$ any choice of Bayesian estimator is possible, such the Maximum A Posteriori [3] or the Minimum Mean Square Error Estimate, i.e. the conditional expectation. We propose using the latter to achieve an unbiased estimate, hence $\hat{x}_i$ is given by

$$\hat{x}_i = \mathrm{E}[x_i|Y_i] = \frac{\alpha_{i|i}}{\beta_{i|i}} \quad (32)$$

As previously mentioned, the filter proposed in [6] is biased, and we can use their same approach to show the unbiased nature of the proposed method. Supposed the state is static $x_i = x_0 \ \forall i$ and that N i.i.d. measurements are applied recursively, then

$$\alpha_{0|i} = \alpha_0 + \sum_{i=1}^{N} \alpha_w = \alpha_0 + N\alpha_w \quad (33)$$

$$\beta_{0|i} = \beta_0 + \sum_{i=1}^{N} \frac{\beta_w}{y_i} = \beta_0 + \beta_w \sum_{i=1}^{N} \frac{1}{y_i} \quad (34)$$

Taking the limit as $N \to \infty$ and noticing that $y_i = w_i x_0$

$$\alpha_{0|i} \to N\alpha_w \text{ as } N \to \infty \quad (35)$$

$$\beta_{0|i} \to \frac{\beta_w}{x_0} \sum_{i=1}^{N} \frac{1}{w_i} = \frac{\beta_w}{x_0} \frac{N\alpha_w}{\beta_w} = \frac{N\alpha_w}{x_0} \text{ as } N \to \infty \quad (36)$$

Finally, the estimate is unbiased since it averages to the true state:

$$\hat{x}_i = \frac{\alpha_{i|i}}{\beta_{i|i}} = x_0 \text{ as } N \to \infty$$

IV. EXAMPLE

We consider a random sequence x_i which is initial distributed according to $x_0 \sim \mathscr{G}(x_0; \alpha_0, \beta_0)$. The variable evolves deterministically through time as shown in Equation (37).

$$x_i = c x_{i-1} \quad (37)$$

Provided $c > 0$, it can be shown that the predicted distribution of $x_{i|i-1}$ from the distribution of $x_{i-1|i-1}$ is given by

$$p_{x_i}(x_i|Y_{i-1}) = \mathscr{G}\left(x_i; \alpha_{i-1|i-1}, \beta_{i-1|i-1}/c\right). \quad (38)$$

At each time-step, we receive a measurement, y_i, which has been corrupted by inverse gamma multiplicative noise, $w_i \sim \mathscr{G}^{-1}(x_i; \alpha_{w_i}, \beta_{w_i})$, as shown in Equation (39).

$$y_i = w_i x_i \quad (39)$$

The likelihood, $L(x_i|y_i)$, has been derived previously and can be found according to Equation (25). Through application of Bayes Theorem, it can be shown that the posterior, $p_{x_i}(x_i|Y_i)$, is proportional to

$$p_{x_i}(x_i|Y_i) \propto x_i^{\alpha_w+\alpha_{i-1|i-1}-1} e^{-x_i\left(\frac{\beta_w}{y_i}+\beta_{i-1|i-1}\right)}. \quad (40)$$

Since all other terms are constant with respect to the random variables and all probability density functions integrate to one, we can conclude that the posterior is a gamma distribution defined as

$$p_{x_i}(x_i|Y_i) = \mathscr{G}(x_i; \alpha_w + \alpha_{i-1|i-1}, \frac{\beta_w}{y_i} + \beta_{i-1|i-1}). \quad (41)$$

With these results, we are able to exactly represent a recursive Bayesian estimator. Evaluation of the conditional mean during either the prediction or update phase of the recursion results in an unbiased prediction or estimate of x_i.

To demonstrate this fact, we use a Monte Carlo simulation. We draw N_{MC} i.i.d. samples of x_0 and propagate them through time from $i = 0$ to $i = 10$ according to Equation (37). At each time-step, we draw independent noisy measurements according to Equation (39). All paremeters used to define the distributions and propagation are shown in Table I.

TABLE I
MONTE CARLO PARAMETERS

Parameter	Value
N_{MC}	1E6
α_0	10.0
β_0	10.0
α_w	22.0
β_w	21.0
c	1.10

A Bayesian estimator is initialized based on the distribution of x_0. Each measurement is used to generate a distribution of the state estimate at the current time-step according to Equation (41). A numerical estimate of the state, $\hat{x}_i$, can be found by calculating the expectation of this distribution according to Equation (2). Likewise, the estimated variance, $\hat{\sigma}_i^2$, was found with Equation (3). The distribution is transformed into a prediction at the next time step according to Equation (38).

At each time-step, the estimate error can be found according to Equation (42). The mean and variance of the errors across all Monte Carlo runs for each time-step are calculated. From the variance positive and negative 3σ deviation bounds can be found. Figure 1 shows the error at each time step for all Monte Carlo runs. Also plotted are the mean error and $\pm 3\sigma$ bounds along with the mean estimated $\pm 3\sigma$ bounds which are derived from the estimated variance, $\hat{\sigma}_i^2$.

$$e_i = x_i - \hat{x}_i \tag{42}$$

Figure 1 demonstrates that the Bayesian update derived in this work is unbiased for all time-steps (estimation error is zero mean). Furthermore, the error variance and mean estimated variance overlap, indicating a consistent filter which accurately predicts its own performance.

To further evaluate filter performance, the performance of an Extended Kalman Filter (EKF) applied to this same problem is shown in Figure 2. The estimated variance found by the EKF is very close to the estimated variance of our method due to the fact that the EKF essentially performs moment matching on the mean and variance of the posterior in the update phase. However, the actual performance of the EKF is worse than its predicted performance. Hence the EKF is an overconfident inconsistent estimator.

Table II shows the accuracy of the two methods in terms of Root Sum Squared (RSS) error.

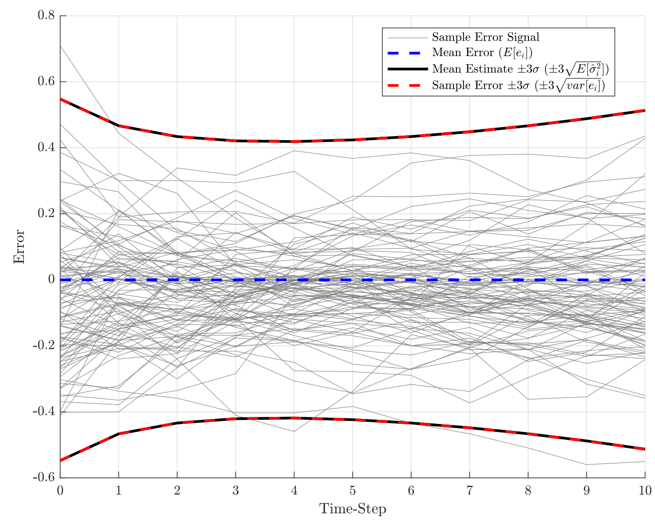

Fig. 1. Demonstration of filter performance during Monte Carlo simulation. Filter is shown to be unbiased and consistent. Note that estimate 3σ grows with time due to system dynamics.

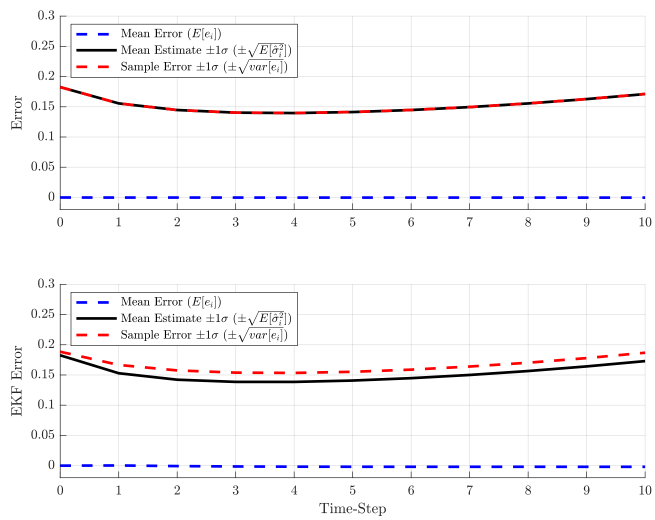

Fig. 2. Comparison between the recursive gamma filter (top) and an EKF (bottom). Note that 1σ values of the EKF's estimate match the 1σ values of the gamma filter. However, the EKF is inconsistent.

TABLE II
FILTERS PERFORMANCE

RSS Error of Gamma Filter:	0.5109
RSS Error of EKF:	0.5542

V. CONCLUSION

We have presented a recursive Bayesian estimator for Gamma-distributed random variables. We have theoretically proven this estimator is unbiased for the static case, and shown it is unbiased and consistent through Monte Carlo for the dynamic case. In addition, we have also shown clear performance benefits compared to the EKF for the dynamic case.

In the course of our derivation we have made several assumptions about the form of the system dynamics and measurement function. These assumptions can likely be relaxed to include additional models. In particular, all that is required for the update is that the predicted distribution is a conjugate prior to the likelihood. The gamma distribution is known as being conjugate prior to several distributions. This fact can be used to relax assumptions on the form of the measurement function and the measurement noise. These topics will be explored in a future work.

REFERENCES

[1] J. Engel, T. Schöps, and D. Cremers, "Lsd-slam: Large-scale direct monocular slam," in *European conference on computer vision*. Springer, 2014, pp. 834–849.

[2] C. Marcus and R. Zanetti, "Direct visual slam with a filter-based lidar measurement incorporation," in *31st AAS/AIAA Space Flight Mechanics Meeting*. AAS/AIAA, 2021.

[3] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with applications to tracking and navigation: theory algorithms and software*. John Wiley & Sons, 2004.

[4] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME–Journal of Basic Engineering*, vol. 82, no. Series D, pp. 35–45, 1960.

[5] H. W. Sorenson, *Kalman filtering: theory and application*. IEEE, 1985.

[6] F. Govaers and H. Alqaderi, "A gamma filter for positive parameter estimation," in *2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*. IEEE, 2020, pp. 40–45.

[7] F. Louzada and P. L. Ramos, "Efficient closed-form maximum a posteriori estimators for the gamma distribution," *Journal of Statistical Computation and Simulation*, vol. 88, no. 6, pp. 1134–1146, 2018.

[8] F. Louzada, P. L. Ramos, and E. Ramos, "A note on bias of closed-form estimators for the gamma distribution derived from likelihood equations," *The American Statistician*, vol. 73, no. 2, pp. 195–199, 2019.

[9] "Transformations of Random Variables," 2 2021, [Online; accessed 2021-08-30]. [Online]. Available: https://stats.libretexts.org/@go/page/10147